American Journal of Epidemiology
Vol. 132, No. 4
Printed in U.S.A.

A Brief Original Contribution

DOES NONDIFFERENTIAL MISCLASSIFICATION OF EXPOSURE ALWAYS BIAS A TRUE EFFECT TOWARD THE NULL VALUE?

MUSTAFA DOSEMECI, SHOLOM WACHOLDER, AND JAY H. LUBIN

Dosemeci, M. (National Cancer Institute, NIH, Rockville, MD 20892), S. Wacholder, and J. H. Lubin. Does nondifferential misclassification of exposure always bias a true effect toward the null value? *Am J Epidemiol* 1990;132:746-8.

The authors present some examples to demonstrate that in certain nondifferential misclassification conditions with polychotomous exposure variables, estimates of odds ratios for categories at intermediate level of risk can be biased away from the null or can change direction. In addition, the authors present two examples to demonstrate that the slope of the dose-response trend for the true distributions can change direction, creating a false inverse trend, even if the misclassification is nondifferential.

bias; case-control studies; dose-response trend; epidemiologic methods; misclassification, nondifferential

It is often stated that nondifferential misclassification of exposure, which is independent of disease status, can only bias an estimate of a true positive odds ratio downward and not away from or beyond the null value (1–6). Although this is true for dichotomous exposures, we present two examples to demonstrate that the bias is not necessarily downward when a polychotomous exposure measure is used. We further show that an estimate of trend in an ordered polychotomous exposure can change direction in the presence of nondifferential misclassification.

Consider the correctly classified data from a hypothetical case-control study shown in "a" in table 1. Assume that for both cases and controls, 40 percent of subjects truly in the high exposure group are incorrectly assigned to the low exposure group, yielding the misclassified distribution shown in "b" in table 1. In this example, the estimate of the odds ratios for the high exposure category did not change, but contrary to our previous understanding, the odds ratio for the low exposure category is elevated from 2.00 to 3.14.

The second example involves nondifferential misclassification between two nonadjacent exposure categories ("c" in table 1). Here, 40 percent of the subjects in the high exposure category are incorrectly assigned to the no exposure category. As expected, the odds ratio for the high exposure category is reduced toward the null; however, for the low exposure category, the odds ratio is reduced below the null, indicating a protective effect, when in reality, the low exposure group is associated with elevated risk.

In addition, we present two examples to demonstrate that, under some nondifferential misclassification conditions, it is possible to create a false negative dose-response trend in the odds ratios, when the true trend is positive. Shown as "a" in table

Received for publication February 14, 1990, and in final form April 30, 1990.

From the Epidemiology and Biostatistics Program, National Cancer Institute, Rockville, MD.

Reprint requests to Dr. Mustafa Dosemeci, Occupational Studies Section, Epidemiology and Biostatistics Branch, National Cancer Institute, NIH, Building EPN, Room 418, Rockville, MD 20892

2 are the results from a hypothetical case-control study showing a positive dose-response trend. Assume that, for both cases and controls, all exposures are categorized correctly except that 40 percent of subjects who are truly in the high exposure group are misclassified into the no exposure group, and 40 percent from the low exposure group are misclassified into the high exposure group, resulting in the observed data shown in "b" in table 2. While the odds ratios in the original data are 2.0 and 6.0 for the low and high exposure categories, respectively, they are now 0.5 and 0.5, respectively, in the misclassified table. With exposure scores of 0, 1, and 2 for none, low, and high exposure categories, respectively, the estimated slope for the correctly classified data is +1.05, while it is −0.03 for the misclassified data. In the second example, 60 percent of subjects who are truly in the no exposure group are misclassified into the high exposure group and 60 percent from the high exposure group are misclassified into the no exposure group, changing the odds ratios for the low and high exposure categories from 2.0 and 8.0 to 0.9 and 0.7, respectively, with a false negative dose-response trend of −0.16 in the misclassified table. Note that it can be shown analytically that the false inverse trend cannot be created under any nondifferential misclassification conditions when the true distribution has no dose-response trend.

TABLE 1

Examples of the effects of nondifferential misclassification on risk estimates

Disease status	Exposure status		
	None	Low	High
a. Reference distribution			
Cases	100	200	600
Controls	100	100	100
Odds ratios	1.00	2.00	6.00
b. Misclassified between adjacent categories			
Cases	100	440	360
Controls	100	140	60
Odds ratios	1.00	3.14	6.00
c. Misclassified between nonadjacent categories			
Cases	340	200	360
Controls	140	100	60
Odds ratios	1.00	0.82	2.47

DISCUSSION

Our examples demonstrate that in some situation with particular forms of nondifferential misclassification, estimates of odds ratios for categories at intermediate levels of risk can be biased away from or beyond the null. The possibility of positive bias in the intermediate exposure odds estimates was pointed out by Walker et al., who claimed, in their work on proxy respondents, "It is not possible in the general

TABLE 2

Examples for the creation of a false negative dose-response trend for odds ratios in a hypothetical case-control study, when nondifferential misclassification occurs among nonadjacent exposure categories

Disease status	Exposure status					
	Example I			Example II		
	None	Low	High	None	Low	High
a. Reference distributions						
Cases	4	800	120	53	40	60
Controls	4	400	20	424	160	60
Odds ratios	1.00	2.00	6.00	1.00	2.00	8.00
b. Misclassified distribution						
	(40% misclassification)			(60% misclassification)		
Cases	52	480	392	57	40	56
Controls	12	240	172	206	160	278
Odds ratios	1.00	0.46	0.53	1.00	0.90	0.73

case to predict the direction of bias that results from [nondifferential] misclassification" (7, p. 907). We also showed that the direction of the estimate of trend can be reversed. These examples indicate that surprising distortions can arise from nondifferential misclassification.

We are aware of the fact that the misclassification patterns presented in the examples are more extreme than those found in most epidemiologic studies. It can be, however, difficult to rule out the kinds of misclassification patterns needed to cause these distortions in some occupational and nutritional studies. For example, four different solvents, stoddard solvent, carbon tetrachloride, trichloroethylene, and perchloroethylene have been the major solvents for dry cleaning operations at various periods since the beginning of the century. In retrospective assessment of exposure to trichloroethylene, it is quite possible that highly exposed dry cleaners could be incorrectly classified as nonexposed, and vice versa, if the timing of the switchovers to and from trichloroethylene were determined incorrectly.

While we do not wish to suggest that the problem we have identified is commonplace, caution is needed in interpreting results in the presence of misclassification, even if the misclassification is known to be nondifferential.

References

1. Copeland KT, Checkoway H, McMichael AJ, et al. Bias due to misclassification in the estimation of relative risk. Am J Epidemiol 1977;105:488–95.
2. Marshall JR, Priore R, Graham S, et al. On the distortion of risk estimates in multiple exposure level case-control studies. Am J Epidemiol 1981;113:464–73.
3. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. Belmont, CA: Lifetime Learning Publications, 1982.
4. Blettner M, Wahrendorf J. What does an observed relative risk convey about possible misclassification? Methods Inf Med 1984;23:37–40.
5. Flegal KM, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. Am J Epidemiol 1986;123:736–51.
6. Rothman KJ. Modern epidemiology. Boston: Little, Brown, and Company, 1986.
7. Walker AM, Velema JP, Robins JM. Analysis of case-control data derived in part from proxy respondents. Am J Epidemiol 1988;127:905–14.